SYSTEMS BIOLOGY

Human PBMC scRNA-seq–based aging clocks reveal ribosome to inflammation balance as a single-cell aging hallmark and super longevity

Hongming Zhu[1]†, Jiawei Chen[2]†, Kangping Liu[2]†, Lei Gao[1,3], Haiyan Wu[1], Liangliang Ma[4], Jieru Zhou[4], Zhongmin Liu[1]*, Jing-Dong J. Han[2]*

Quantifying aging rate is important for evaluating age-associated decline and mortality. A blood single-cell RNA sequencing dataset for seven supercentenarians (SCs) was recently generated. Here, we generate a reference 28-sample aging cohort to compute a single-cell level aging clock and to determine the biological age of SCs. Our clock model placed the SCs at a blood biological age to between 80.43 and 102.67 years. Compared to the model-expected aging trajectory, SCs display increased naive CD8+ T cells, decreased cytotoxic CD8+ T cells, memory CD4+ T cells, and megakaryocytes. As the most prominent molecular hallmarks at the single-cell level, SCs contain more cells and cell types with high ribosome level, which is associated with and, according to Bayesian network inference, contributes to a low inflammation state and slow aging of SCs. Inhibiting ribosomal activity or translation in monocytes validates such translation against inflammation balance revealed by our single-cell aging clock.

INTRODUCTION

Aging is a complex process of functional decline of an organism, which is accompanied by many degenerative diseases. Aging clocks refer to linear or nonlinear models trained on omics data to predict age and aging rate, which, measured by the difference between chronological and predicted age (AgeDiff), can be used to quantify aging at different levels (*1*). For example, human peripheral blood transcriptome clocks have a mean absolute difference (MAD) between chronological age and predicted age of 7.8 years (*2*), while blood proteome–based prediction reaches a Pearson correlation coefficient (PCC) between predicted and actual age of 0.93 to 0.97 (*3*), and DNA methylome has reached MAD of 4.9 years in the whole blood (*4*) and 3.6 years in heterogeneous tissues (*5*). A recent 50 blood cytokine–based immune aging clock has a MAD of 15.2 years (*6*). We have established a three-dimensional human facial image–based linear model with MAD of 6.1 years (*7*) and a deep learning artificial intelligence (AI) model of 2.8 years (*8*). Fast or slow agers (AgeDiff > MAD) display substantial aging-related physiological decline or delay than well predicted individuals (*2–7*). However, how much of the heterogeneity in aging rate is contributed by the heterogeneity at the single-cell level has not been examined. Although a few single-cell RNA sequencing (scRNA-seq) analyses comparing two points of age (young and old) have found many age-related single-cell level changes in human and monkey ovary (*9*, *10*), and various tissues in mice (*11–17*), the two points

data cannot establish an age and aging rate predictor to assess the aging rate or its heterogeneity at the single-cell level. Several single-cell clocks have been developed in mice including multi-tissue single-cell DNA methylation clocks (*18*) and brain single-cell aging clocks (*19*). While 45-sample human peripheral blood mononuclear cell (PBMC) scRNA-seq data across different ages are available (*20*), the quality and sparsity of the data do not allow reliable separation of known cell types or clear mapping of cell identities, in addition to the lack of precise age information.

A very small group of individuals, the supercentenarians (SCs), >110 years old, not only achieve super longevity but also have a disease-free healthy and active life even at very old age. Recently, a precious blood scRNA-seq for 7 SCs was generated (*21*); however, without a quantitative aging clock model, there is no proper control or reference to decipher SCs' cellular and molecular features associated with their extreme longevity, as neither young nor old samples can be rationally used for comparison. Specifically, due to the lack of a natural aging cohort to establish single-cell aging clock, it was not possible to examine what single-cell aging features were delayed or reversed in these SCs, and what were the biological ages of the SCs compared to the common single-cell aging trajectories, or what allow them move into a different aging trajectory. To address these problems, here we generated a 17-sample human PBMC scRNA-seq dataset from volunteers in Shanghai, China, evenly distributed from age 28 to 77 years, which we termed the Shanghai East Hospital natural aging cohort (or SE cohort for short). On the basis of the cell type composition changes in SE cohort and another cohort aged from 72 to 100 years (*22*), we built a linear aging clock, then applied to SCs, and revealed the cellular and molecular hallmarks deviating SCs' actual age from the model-predicted age at the single-cell and cell type level (Fig. 1A).

[1]Translational Medical Center for Stem Cell Therapy & Institute for Regenerative Medicine, Shanghai East Hospital, Tongji University School of Medicine, Shanghai 200120, P.R. China. [2]Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Center for Quantitative Biology (CQB), Peking University, Beijing 100871, P.R. China. [3]Division of Life Sciences and Medicine, University of Science and Technology of China (USTC), Hefei, Anhui 230001, P.R. China. [4]Department of Health Management, Shanghai East Hospital, Tongji University School of Medicine, Shanghai 200120, P.R. China.
*Corresponding author. Email: jackie.han@pku.edu.cn (J.-D.J.H.); liu.zhongmin@tongji.edu.cn (Z.L.)
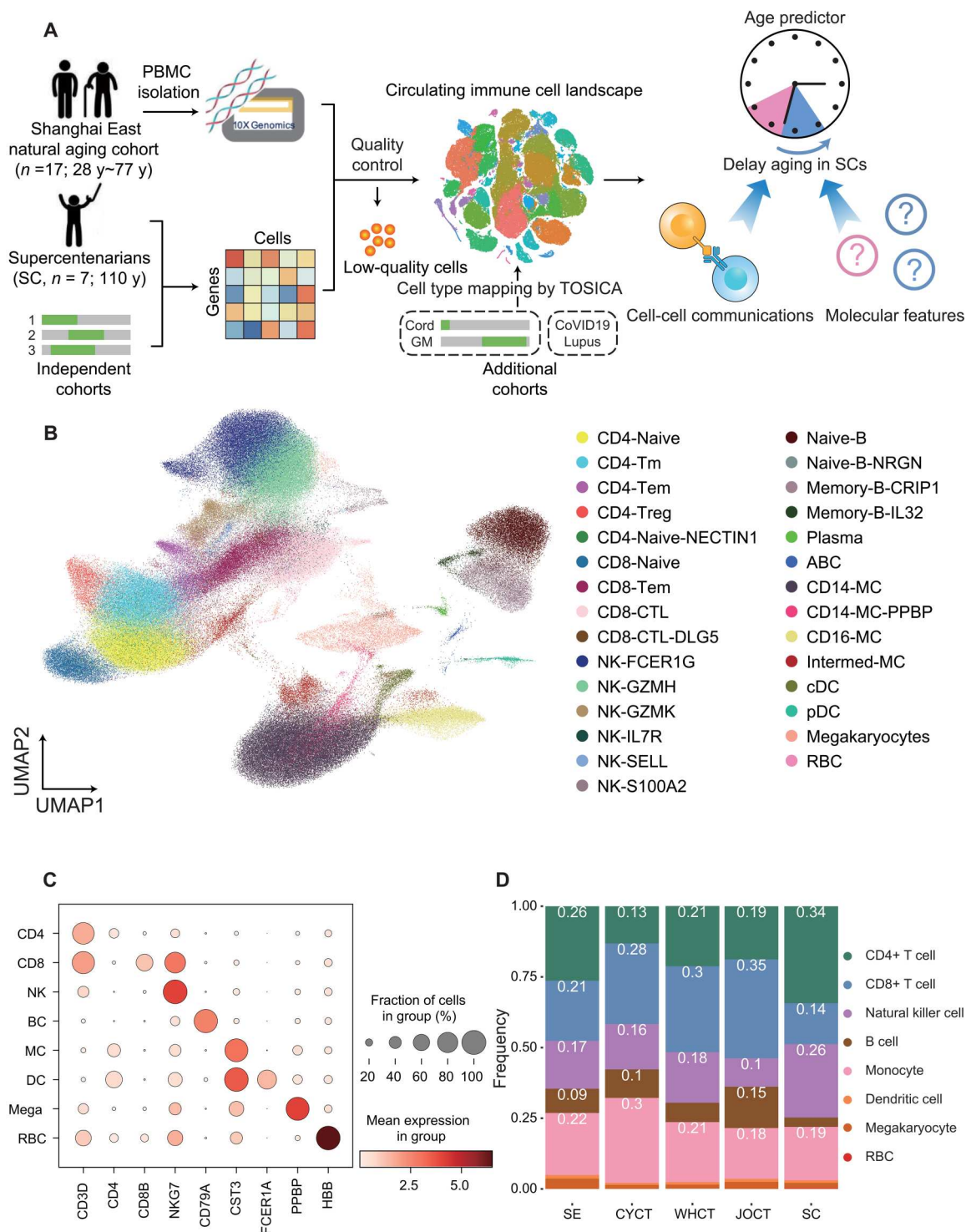†These authors contributed equally to this work.

Fig. 1. Single-cell RNA sequencing (scRNA-seq) landscape of human blood peripheral blood mononuclear cell (PBMC) during aging. (**A**) Study design and flow chart of the human blood PBMC scRNA-seq aging clock analysis. The scRNA-seq data from SE, Chinese young cohort (CYCT), Wuhan cohort (WHCT), Japanese old cohort (JOCT), and supercentenarian (SC) cohort are merged to classify cell types, and then cell type annotations are transferred to additional cohorts by TOSICA for further analysis. The green bars show the age range of each cohort. (**B**) Two-dimensional uniform manifold approximation and projection (UMAP) visualization of scRNA-seq data from PBMCs of all cohorts. Each point is a cell; colors are based on subtype annotation defined by the Louvain clustering algorithm and marker genes. (**C**) Expression levels of significantly differentially expressed genes for eight major cell types. The color and size of each dot represent the expression level and cell fraction of the marker genes, respectively. (**D**) Proportion of cell types shown in (C), in different cohorts. Cell type frequency is labeled in white text.

RESULTS

scRNA-seq reveals cell type composition in aging cohorts

To build a single-cell transcriptome–based aging rate predictor, we collected the PBMCs from 17 individuals (6 males and 11 females with similar age distributions) from 28 to 77 years of age, which we termed the Shanghai East Hospital natural aging cohort (or SE cohort for short), and subjected them to 10X Genomics scRNA-seq. For each individual, an average of 8545 cells were sequenced and analyzed with an average of 110,364 reads per cell (Fig. 1A). We processed these data together with published PBMC scRNA-seq datasets for further analysis. These include two Chinese cohorts, Chinese young cohort (CYCT) (23) and Wuhan cohort (WHCT) (24), and two Japanese cohorts, Japanese old cohort (JOCT) and SCs (21). Each of the cohorts contains five individuals. After removing 25,111 cells due to poor quality (cells with >10% mitochondrial gene expression or <200 genes detected), a total of 131,972 single cells (average of 7763 cells per sample) are retained for the SE datasets and a total of 23,796 cells for CYCT, 45,923 cells for WHCT, 19,252 cells for JOCT, and 38,417 cells for SC (fig. S1A). We then removed batch effect between SE datasets and other independent cohorts, and identified red blood cells (RBCs), megakaryocytes (Mega), and six major immune cell lineages—CD4$^+$ T cells (TC), CD8$^+$ T cells, natural killer cells (NK), monocytes (MC), and dendritic cells (DC), based on the expression of canonical lineage markers and cluster-specific marker genes, and visualized them by Uniform Manifold Approximation and Projection (UMAP) (Fig. 1B). The presence of cells from both SE datasets and other independent datasets in each cluster demonstrates that all clusters contained cells from multiple samples and thus are not segregated due to dataset-specific batch effects (fig. S1B).

Immune cell subtypes are dissected by recluster

To separate immune cell subtypes, we separately reclustered the cells of NK and TC lineage, BC lineage, and MC and DC lineage. Using the most significantly up-regulated genes in each cluster, we identified 29 cell subtypes in total. The CD4$^+$ T cells were subdivided into five classes: CCR7^{high} naive CD4$^+$ T cells (CD4 Naive), CCR7^{high} NECTIN1^{high} naive CD4$^+$ T cells (CD4 Naive-NECTIN1), CCR7^{med} CCR6$^-$ memory CD4$^+$ T cells (CD4 Tm), CCR6$^+$ effector memory CD4$^+$ T cells (CD4 Tem), and FOXP3$^+$ regulatory T cells (CD4 T$_{reg}$) (Fig. 1B and fig. S1, C and D). The CD8$^+$ T cells were subdivided into four classes: CCR7$^+$ naive CD8$^+$ T cells (CD8 Naive), GZMK$^+$ effector memory CD8$^+$ T cells (CD8 Tem), GZMB$^+$ GNLY$^+$ cytotoxic CD8$^+$ TCs (CD8 CTL), and DLG5^{high} cytotoxic CD8$^+$ TCs (CD8 CTL-DLG5) (Fig. 1B and fig. S1, C and D). The NKs were subdivided into six subtypes by the most significantly up-regulated genes in each cluster (fig. S1C). In addition, we identified six major B cell subsets: two groups of IL4R$^+$ IGHD$^+$ naive B cells (Naive B and Naive B-NRGN); two groups of CD27$^+$ IGHG1$^+$ memory B cells (Memory B-IL32 and Memory B-CRIP1); plasma cells or so-called antibody-secreting cells, expressing high level of immunoglobulin genes MZB1; and a subset of ITGAX$^+$ B cells defined as age-associated B cells (ABCs) (Fig. 1B and fig. S1, E and F). As for myeloid cells (including MCs and DCs), we identified six transcriptionally distinct subsets: CD14^{high} CD16$^-$ classical monocytes (CD14-MCs), CD14^{high} CD16$^-$ PPBP^{high} classical monocytes (CD14-MC-PPBP), CD14$^{+/-}$ CD16^{high} nonclassical monocytes

(CD16-MCs), CD14$^+$ CD16$^{+/-}$ intermediate monocytes (Intermed-MCs) (fig. S1, G and H), CD1c$^+$ conventional DCs (cDCs), and CLEC4C$^+$ plasmacytoid DCs (pDCs) (fig. S1, G and H). These subtypes of cells were then used for downstream analyses.

Cell type proportion-based clock models reveal SCs to be much younger than their actual age

With increased age, various cell types change in proportions (Fig. 2A). We therefore tested using cell type proportions of the scRNA-seq data to build a "clock" to predict age. We first trained a partial least square regression (PLSR) linear model and used leave-one-out cross-validation (LOOCV) to assess the accuracy of the model. To compare prediction power across models and to prevent overfitting, all PLSR models in this study included only the top three components. We then calculated variable importance in projection (VIP) to rank the overall contribution of each variable to the PLSR model, and the variables with VIP score > 1 are considered to be high explanatory variables (25). These include NK-GZMH, CD8-CTL, CD8-Naive, CD4-Tm, CD4-Naive, Naive-B, NK-FCER1G, CD14-MC, and Memory-B-CRIP1 contributing significantly to the PLSR model (Fig. 2B). To extend the age spectrum of training data, using TOSICA (26), we further mapped the cell types of a published PBMC dataset generated by Guangdong Medical University that contains very old (72 to 100 years) samples (GM) (22) (fig. S2A) and trained a PLSR model. Although the GM dataset contains frail and old control samples, frail samples were reported to have elevated expression of the NEAT1 long noncoding RNA (lncRNA) in a fraction of CD14 monocytes (22), which are not a cell type contributing to our cell type composition–based aging clocks, and we do not find that they differ significantly in cell type compositions (fig. S2B) or in scRNA-seq clock-predicted cAgeDiff (fig. S2C); therefore, we include both set of samples as expanded training samples. On the basis of this extended dataset (SE + GM), we find that the three-component PLSR model based on only the top 7 cell types generates a model with PCC between predicted and chronological age of 0.88 and MAD of 8.36 (Fig. 2C). Through saturation analysis, we find that the 28 reference samples we used for clock building are sufficient to reach a stable high accuracy (fig. S2D). We also used the cell type proportions to train an elastic net (EN) regression linear model and used the weight to rank the overall contribution of each variable to the EN model (fig. S2E). The EN model based on the top 7 cell types had a similarly accurate predictive power than PLSR model, with PCC and MAD of 0.89 and 7.86, respectively, between predicted and chronological age in SE + GM cohorts (fig. S2F), whereas the accuracy of PLSR model is more similar between training and independent cohorts, suggestive of a slightly better generalizability. Five of the top 7 cell types overlap between PLSR and EN model (fig. S2G), hinting at their critical importance at age prediction by blood cell type compositions.

To further validate our models, we tested them on blood scRNA-seq data of three independent healthy cohorts, five CYCT individuals with an age range from 29 to 58 years old, five WHCT individuals from 37 to 71 years, and five JOCT individuals with an age range from 50s to 80s (age accurate to decades instead of years; thus, an error exceeding 5 years is expected by definition). Our seven-cell type PLSR model had MAD of 10.06, 10.91, and 10.16 and PCC of 0.43, 0.69, and 0.28 to chronological age in these three cohorts, respectively (fig. S2H). Similar MADs were found with the

Fig. 2. Single-cell RNA sequencing (scRNA-seq)–based single-cell composition aging clock and age delay of supercentenarians (SCs). (**A**) Spearman rank correlation coefficient (RCC) of age to proportion of cell types with RCC > 0.25 across all SE samples. (**B**) Variable importance in projection (VIP) values of cell types with VIP ≥ 1 in partial least square regression (PLSR) clock model. (**C**) PLSR clock model trained and cross-validated on cell type proportions in SE + Guangdong Medical University (GM) samples. Each dot represents one individual, colored by cohort. (**D**) Chronological age and age-corrected PLSR (based on 28 SE + GM samples) predicted age of independent cohorts. Each dot represents one individual. (**E**) PLSR (based on 28 SE + GM samples)–predicted age-corrected difference between chronological and predicted age (cAgeDiff) of individuals with mild and severe COVID-19 compared to age-matched healthy controls. Student's t test P value is shown on the top. (**F**) PLSR (based on 28 SE + GM samples)–predicted cAgeDiff of COVID-19–positive patients at convalescence and progression stages. (**G**) PLSR (based on 28 SE + GM samples)–predicted cAgeDiff of systemic lupus erythematosus (SLE) patients of managed, flare, and treated states compared with age-matched healthy controls. Each point represents one individual, and the error bar represents mean and SD. Student's t test P values are shown on the top. (**H**) Anti-correlation between the proportions of CD8-CTL and CD8-Naive cells in CD8+ T cells across SE + GM and SC samples. Pearson's correlation coefficient and P value are shown.
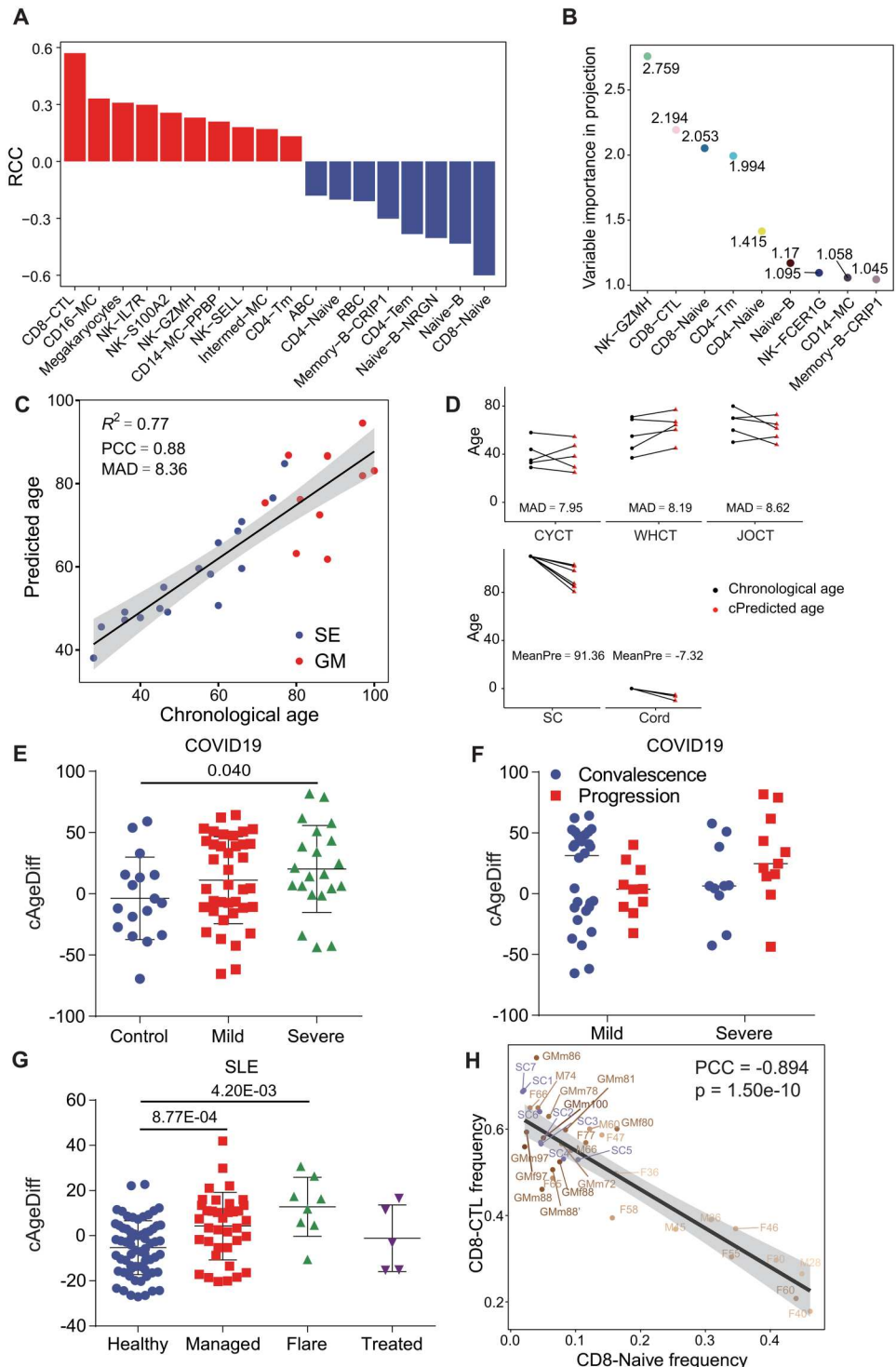
seven–cell type EN model (fig. S2I). When plotting the difference between predicted and chronological age (AgeDiff), we noticed that there is a slight systemic underprediction (negative AgeDiff) in old age and an overprediction (positive AgeDiff) in young age in our PLSR and EN models just as in other clock models (*8*); thus, we further corrected this systemic bias using Loess model to derive age-corrected AgeDiff (cAgeDiff) as described previously (*8*)

(fig. S3A). After this correction, the MADs in independent control cohorts are 7.95, 8.19, and 8.62 and PCC of 0.64, 0.85, and 0.57 to chronological age. It should be noted that age correction does not change MAD of the model but will enhance PCC to chronological age. Since our single-cell aging clock is linear to age, we expect it to have linear extensibility to extrapolate to either end of the age spectrum that is not included in the training data. More than 28 years

younger than the youngest training sample, TOSICA-mapped umbilical cord blood samples (Cord) are predicted by our clock model to be −10.10 to −5.42 years old (Fig. 2D). Toward the other end of the age spectrum, SCs who are 20 years older than the oldest training sample are predicted to be 80.43 to 102.67 years old, on average 18.64 years younger than their chronological age. The age-corrected cell type EN model placed all SCs to below 110, as young as the average of 95.23 years old, and umbilical cord blood (Cord) to 7.12 to 12.68 years old (fig. S3B). We further confirmed that the age-corrected clock was not affected by batch-related factors, such as sex, datasets, and cell numbers (fig. S3, C to E).

To further test clock performance, we applied our age clocks to other independent disease datasets including a COVID-19 dataset and a systemic lupus erythematosus (SLE) dataset (fig. S2A). Our cell type composition–based PLSR model reveals a higher average biological age of individuals in disease states than age-matched controls as shown by cAgeDiff, which increases with the disease severity (Fig. 2, E to G). Perhaps due to the small sample size, only severe symptom COVID-19 patients are significantly older (Fig. 2E; t test, $P = 0.040$) and SLE symptoms managed and flare patients are highly significantly older than controls (Fig. 2G; t test, $P = 8.77 \times 10^{-4}$ and 4.20×10^{-3}). cAgeDiff predicted by the EN models shows a consistent significant increase in diseases versus controls (fig. S3, F to H). These suggest that both the PLSR and EN cell type composition clock models can identify the increase in immune-related biological age of patients with these diseases.

Consistent with the model-predicted "slow aging" of the SCs (Fig. 2D and fig. S2H), among the seven cell types used in the PLSR and those in EN models, a total of nine cell types (fig. S2G), a highly significant negative correlation exists between two CD8$^+$ T subtypes, CD8$^+$ CTL and CD8$^+$ naive T cells, which significantly increase and decrease with age, respectively, in both SE + GM and SC cohorts, but SCs largely overlap with the 70- to 100-year-old SE + GMs (Fig. 2H). This suggests that the polarization of CD8$^+$ naive T toward CD8$^+$ CTL is an aging hallmark at the blood cell population level.

Age-related cell composition changes are delayed in SCs

Now that we have natural aging cohort-derived clock models, instead of artificially using any particular age group as controls, we use the cellular and molecular deviation of model-predicted versus their actual status in SCs to decipher features associated with their extreme longevity (Fig. 3A). We reverse-calculated the cell type compositions at age 110 years by PLSR clock models (Materials and Methods). Alternatively, based on the top 7 cell types in the PLSR and EN clock models (Fig. 2E), we defined the cell types whose proportions are significantly positively or negatively correlated with age by linear regression model (LR) (with RCC to age slope, $P < 0.05$), without age-related systemic bias (fig. S4A). Then, the expected cell proportions at age 110 are calculated based on the PLSR model or the LR model, and if they are significantly higher or lower than the real proportions in SCs (z score, $P < 0.05$), the cell types are considered as "SC delayed age-up cell types" or "SC delayed age-down cell types" (Fig. 3A). We find that SCs show significantly delayed age-dependent decrease in CD8$^+$ naive T cells, Naive B cells, Memory-B, and NK-GZMK (z score, $P < 0.05$), marginally significantly delayed age-dependent increase in CD4$^+$ Tm (z score, $P < 0.1$) compared to the PLSR model (fig. S4B), and delayed age-dependent decrease in CD8$^+$ naive T cells and delayed age-dependent

increase in CD4$^+$ Tm and Naive B compared to the LR model (fig. S4C). In addition, the frequencies of cell types that contribute to the aging clocks and are delayed in SCs show no differences between the predicted values and real values of individuals from both training and independent cohorts, except in Cord and SC cohorts (fig. S4, D and E). When the proportions of age down-regulated CD8-Naive cells and age up-regulated CD4 Tm cells are plotted across age, there are clear delayed deviations of SCs from the linear trend line over the 28 samples of 28 to 100 years (fig. S4F).

Age-related single-cell gene expression changes are delayed in SCs

We used the sum of gene expressions of all cells belonging to each individual to train PLSR models as pseudo-bulk transcriptome clock model, which predicts the 28 training samples at MAD of 5.04 and PCC of 0.97 to chronological age (Fig. 3B) and predicts SCs to a mean age of 82.49 years (fig. S4G). We used the sum of gene expressions of all cells belonging to each cell type for each individual to train PLSR models as the pseudo-bulk transcriptome clock for the cell type separately with MAD from 4.80 (ABC) to 9.60 (Plasma) years (Fig. 3C). These clocks, more dramatically than the cell type proportion-based clocks, predicted the SC median transcriptome ages to 51.62 to 63.19 years before age correction (fig. S4H) to 65.01 to 75.83 after age correction (Fig. 3C). Then, based on these PLSR clock models, or alternatively defining top contributing genes with VIP > 1 (table S1) as age-related, and predicted their expression levels at 110 using LR models in each cell type, we then defined "SC delayed age-up genes" (up-regulated with age and delayed in SCs) and "SC delayed age-down genes" (down-regulated with age and delayed in SCs) similar to cell type analysis described above (Fig. 3A). Using the LR models, which show no systematic bias against age (fig. S5, A and B, and table S2), we find 2749 and 495 age-up and age-down genes, respectively (table S3), and 1023 age-up and 149 age-down genes significantly delayed in SCs. As validation of the model, we find no significant differences between real and predicted gene expressions of genes that contribute to the aging clocks and are delayed in SCs in the training cohort, as well as in other independent cohorts, except in SC cohorts (fig. S5C). Transcription factor (TF) target enrichment analysis shows that age-up and age-down genes are enriched for targets of FOXR2, ZNF318, PSMB5, PER1, ZZZ3, and NERF (fig. S5, D and F). Among them, FOXR2, ZNF318, PSMB5, and ZZZ3 targets are enriched in SC delayed age-down genes and nuclear factor κB (NFκB) targets are enriched in SC delayed age-up genes (fig. S5, F and G). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses reveal that age-up genes enrich for NFκB/ tumor necrosis factor (TNF)/TLR/NOD/FoxO signaling pathway, virus infection, apoptosis, T cell receptor signaling pathway, cytokine, and PD-1 pathway (fig. S6, A and B), and among them, NFκB/TNF/TLR/NOD/FoxO signaling pathway, virus infection, apoptosis, cytokine, and mitogen-activated protein kinase (MAPK) pathway are significantly delayed in SCs (Fig. 3D), while age-down genes most commonly enrich for ribosome genes (fig. S6, C and D), which are significantly delayed in most cell types of SCs (Fig. 3E and table S4).

Compared to the PLSR model–expected values, there are 111 age-up genes significantly delayed in the SCs, which are related to MAPK signaling pathway, rapamycin (mTOR) signaling pathway, NFκB signaling pathway, and virus infection (fig. S6E), whereas
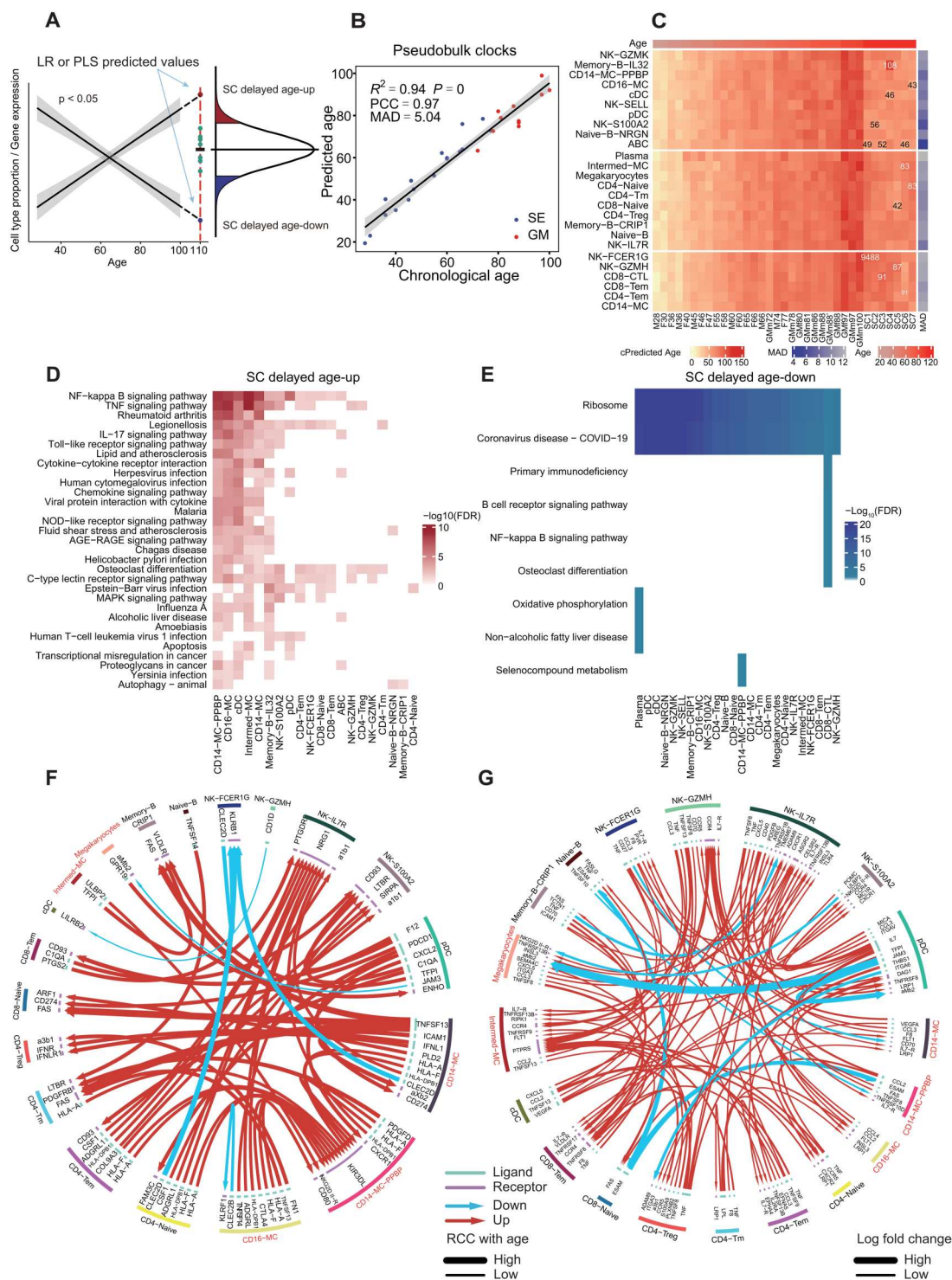
Fig. 3. Aging-related cell type proportion and gene expression changes delayed in supercentenarians (SCs). (**A**) Illustration of the identification and definition of SC delayed age-related changes in cell type proportions and gene expressions. (**B**) Partial least square regression (PLSR) age predictor-trained and cross-validated on pseudobulk transcriptome in SE + GM samples. Each dot represents one individual, colored by cohort. (**C**) Age-corrected PLSR model-predicted age of each SC sample based on the pooled gene expression in each cell type. The maximum and minimum predicted ages of each SC are marked in white and black, respectively. (**D** and **E**) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched in SC delayed age-up and delayed age-down genes in different cell types according to linear regression model (LR) model-predicted gene expression in SCs (Fisher test FDR < 0.01 scaled according to the legend). (**F**) Circos plot showing the top 150 interactions mediated by ligand-receptor pairs between cell types significantly changed with age Pearson correlation coefficient (PCC) *P* < 0.05. The outer ring displays color-coded cell types according to Fig. 1B, and the inner ring represents the ligand-receptor interacting pairs. The line width and arrow width are proportional to *t* test FDR. Colors of edges indicate the changes with age: blue, age-down; red, age-up. (**G**) Circos plot of the SC delayed age-up or age-down interactions between cell types. Blue, SC delayed age-up; red, SC delayed age-down.

1261 age-down genes are significantly delayed in SCs. Again, ribosome- and translation-related genes are enriched in most cell types at the very top of the list (fig. S6F). The delay of these ribosome genes decline in SCs cannot be attributed to batch effect in gene expression because, as a group, average ribosome gene expressions in each cell type do not differ among different cohorts [analysis of variance (ANOVA), $P = 0.53$; fig. S7A], while the age down-regulated ribosome genes identified in SEs also decrease with age in other cohorts (fig. S7B).

Inflammatory cell-cell interactions are repressed in SCs
Altered intercellular communication is an integrative hallmark of aging (27). On the basis of annotated ligand-receptor databases, we calculated the numbers of ligand-receptor pairs in every pair of cell types using CellphoneDB (28). We found that many age-related receptor-ligand pairs are up-regulated during aging among various cell types (PCC, $P < 0.05$). These pairs are mostly related to inflammation, cytokines, and antigen presentation (Fig. 3F). Similar to delayed age-up and age-down cell types and genes, we defined delayed age-up and age-down ligand-receptor pairs in SCs. The age-related down-regulated aMb2 integrin signaling ligand-receptor pairs are generally delayed in SCs for most cell types, and age-up inflammation-related TNF signaling ligand-receptor pairs are also delayed, especially in pDC (Fig. 3G).

Ribosome/translation counteracts inflammation at single-cell level
On the basis of the above results, we wondered whether age-related decline in ribosome levels and age-related increase of inflammation levels are two linked events of the single-cell aging. To quantitatively assess the inflammatory state of each cell, we calculated the inflammatory score for each cell based on inflammatory-related gene list (table S5) as described previously (29). The median expression levels of ribosome genes are used to indicate the level of ribosomes in each cell. In both SE and SC cohort, there is a significant negative correlation (PCC = -0.22, $P < 2.2 \times 10^{-16}$ in SE, PCC = -0.23, $P < 2.2 \times 10^{-16}$ in SC, PCC = -0.30, $P < 2.2 \times 10^{-16}$ combined) between ribosome level and inflammation score across all single cells of all cell types (Fig. 4A). Furthermore, across different cell types, there is also a significant negative correlation between the average ribosome levels and average inflammation scores (RCC = -0.58, $P = 0.0021$ in SE, RCC = -0.59, $P = 0.0020$ in SC, RCC = -0.58, $P = 1.0 \times 10^{-5}$ combined) (Fig. 4B). On the basis of this ribosome level versus inflammation score distribution, cell types can be separated into high-inflammatory and low-ribosome (HI-LR) cell types, which include Intermed-MC, CD14-MC, CD16-MC, CD14-MC-PPBP, and megakaryocytes, and low-inflammatory and high-ribosome (LI-HR) cell types, which are the rest of the cell types (Fig. 4B). In all but one cell type, the total ribosome gene expression level decreases with age in SEs, whereas in HI-LR cell types, SCs show a similar level to young SEs (Fig. 4C). Consistently, principal components analysis (PCA) analysis based on ribosomal and inflammatory gene expressions in each single cell within the same cell type shows that for most cell types, for example, CD4-T$_{reg}$ and CD14-MC, single cells fall into two major cellular states, corresponding to high and low ribosome expression, respectively (fig. S8A). Consistent with the ribosome expression level, the proportion of cells in the high ribosome state decreases with age for most cell types (RCC to age range from -0.64 to 0.43 with a median of -0.23; Fig. 4D and table S6).

Remarkably, SCs consistently have more cells in the high ribosome state than SE in the HI-LR cell types (Fig. 4D). These indicate that HI-LR cell types' high ribosome expression, or more precisely more cells in high ribosome state, is a common feature of super longevity at the single-cell level in the blood.

We also validated the changes of ribosome expression and inflammation scores in COVID-19 and SLE cohorts. Consistent with the balance between ribosome expression and inflammation, with the disease progression in COVID-19 and SLE, the inflammatory score consistently increased and the expression level of ribosomes decreased in all cell types and HI-LR cells, with the changes generally more pronounced in SLE than in COVID-19 (fig. S9).

Furthermore, we find significantly different cell-cell communication between cells in the high ribosome versus in the low ribosome within most cell types (Fig. 4E). In the low-ribosome state, the top 100 ligand-receptor pairs are mostly between CD8-CTL and the rest of the cell types, including MHCII, S100A8/A9-TLR4, CCL5-CCR5, and TNF-related pairs, while these are all significantly lower in high-ribosome state (t test, $P < 0.01$; Fig. 4E and fig. S8B). The most commonly down-regulated ligand-receptor pair in high versus low ribosome state is the interaction between ribosomal protein S19 (RPS19) and its membrane receptor C5AR1 (fig. S8B), which has been reported to suppress immune response (30).

Finally, to understand the causal relationships among ribosome level, inflammation level, age, and aging rate in SE + GM data, we inferred a Bayesian network (BN) based on all single cells in HI-LR and LI-HR cell types, respectively, in the aging rate reference SE + GM data. Here, aging rate is measured by age-corrected AgeDiff, which is the difference between the predicted age and the chronological age based on the cell composition PLSR model and corrected by age. The BN model infers that AgeDiff is dependent on the level of ribosome and inflammation in the HI-LR cell types, and the level of ribosome is in turn dependent on age. The model also infers that a decrease in ribosome level increases in the level of inflammation in the HI-LR cells. This is not the case in LI-HR cell types, where age decreases inflammation score, and inflammation score does not contribute to aging rate (Fig. 4F). These suggest that ribosome abundance, which is often a surrogate of translation and growth rate, may suppress the inflammatory cytokine secretion of the HI-LR cells; in other words, high level of inflammatory cytokine expression might be a consequence of stagnant growth or insufficient basally low ribosomal activities of the HI-LR cell types, and that SCs might benefit from the high cellular ribosome/translation/growth rate to counteract aging-induced chronic inflammation. In contrast, rather than inducing an inflammatory state, the basally low-level inflammatory cytokines might be a consequence of the normal growth of the LI-HR cells (Fig. 4F). Similar relationships can be consistently observed in BN models inferred for each cell type, and based on the interactions in these BNs, the HI-LR cell types automatically aggregate together and segregate from other cell types by unsupervised hierarchical clustering (Fig. 4G).

Finally, to confirm the inferred suppression of inflammation by ribosomal activity (translation) in HI-LR cells, we temporally inhibited translation using a translation inhibitor cycloheximide (CHX), which binds to the 60S ribosomal unit and inhibits translation elongation (31), in human blood primary monocyte cells and THP-1 monocyte cell line, which represents a HI-LR cell type and was
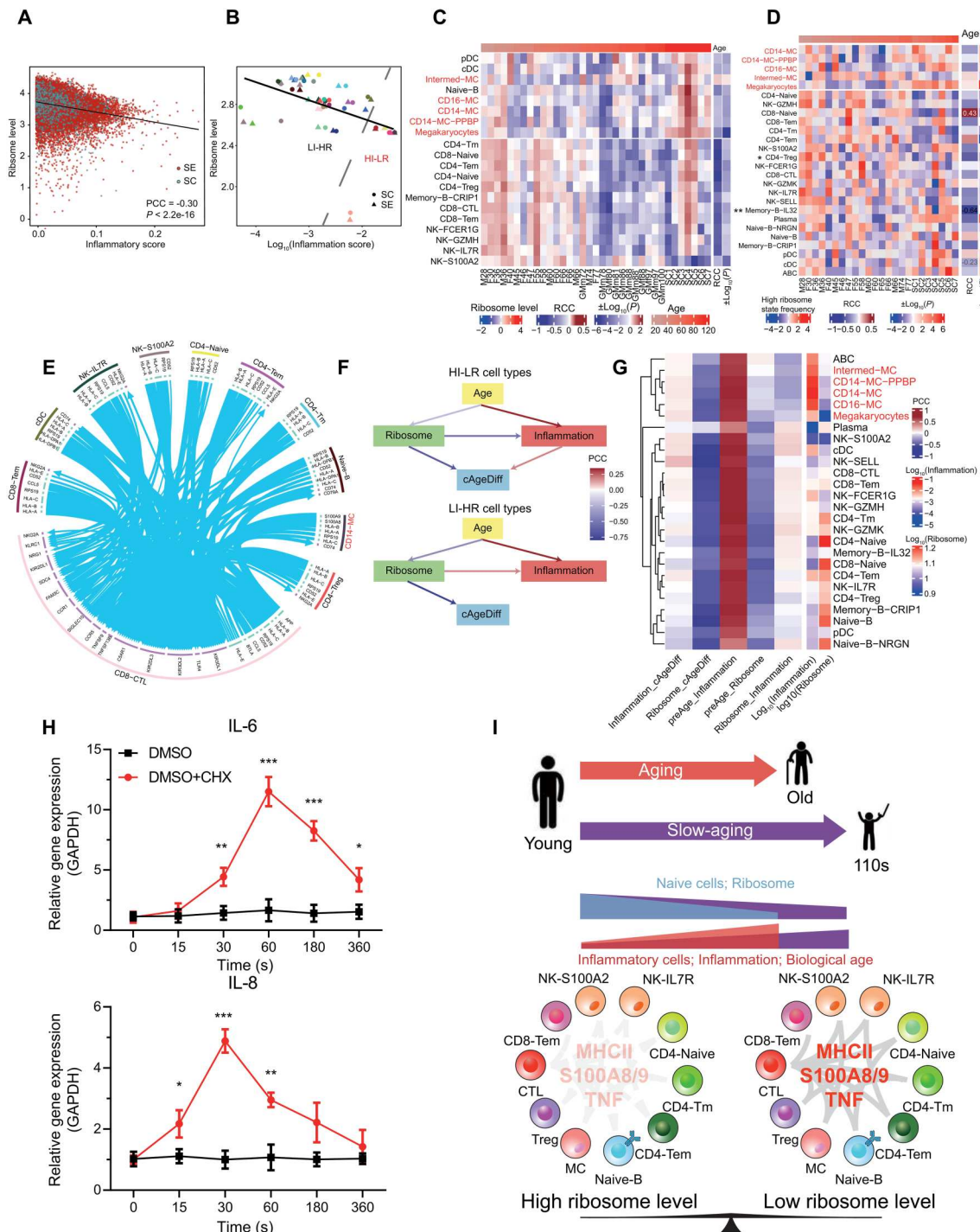
Fig. 4. Relationship between ribosome and inflammation across single cells and cell types during aging. (**A**) Dot plot of ribosome gene expression levels and inflammatory score for each cell in SE and supercentenarian (SC) cohorts. Combined Pearson correlation coefficient (PCC) and *P* value in SE and SC are labeled. (**B**) Dot plot of average ribosome gene expression levels against inflammatory score for all cell types. Combined rank correlation coefficient (RCC) and *P* value in SE and SC are labeled. (**C**) Heatmap of the total ribosome gene expression level in each cell type of each individual. Rank correlation coefficients (RCCs) with age in 28 SE + Guangdong Medical University (GM) samples and *t* test *P* value between SC and predicted value are shown on the right. (**D**) Heatmap of the high ribosome state frequency in each cell type. RCCs with age and *t* test *P* value between SC and SE are shown on the right. The significant RCCs are labeled by *. The maximum, median, and minimum RCCs are marked in white, gray, and black, respectively. (**E**) Circos plot of the down-regulated interactions in high versus low ribosome state (*t* test, *P* < 0.001). (**F**) Bayesian networks inferred among the four factors in HL-LR cell types and LI-HR cell types in 28 SE + GM samples. A directed edge denotes that the occurrence of the target node is dependent on that of the source node. (**G**) Heatmap of the correlation between inflammation score, ribosome levels, age, and age-corrected difference between chronological and predicted age (cAgeDiff) in each cell type of 28 SE + GM samples. The mean level of inflammation and ribosome is labeled on the right. (**H**) Quantitative polymerase chain reaction (PCR) shows the transcript changes of IL-6 and IL-8 in primary monocyte cells treated with vehicle and cycloheximide (CHX) (100 μg/ml) at various time points. **P* < 0.01; ***P* < 0.001 (*t* test). (**I**) Schematic summary of the ribosome and inflammation balance as a hallmark of blood single-cell aging and delayed aging in SCs.

found to secrete senescence-associated cytokines in our previous study (*32*). As expected from the BN model, inhibiting translation caused a significant up-regulation of commonly observed inflammatory cytokines, interleukin-6 (IL-6) and IL-8, in a time-dependent manner in primary monocyte cells (Fig. 4H) and THP-1 (fig. S8C).

DISCUSSION

Together, by generating a reference 28-sample aging cohort, we developed single-cell level single-cell aging clocks based on single-cell type compositions of our reference 17-individual aging cohort, and placed the SCs at a biological age between 80.43 and 102.67 years (Fig. 2D and fig. S2H). Among the independent validation cohorts, both the Chinese WHCT and CYCT and Japanese JOCT are well aligned to their actual ages. Although the single-cell aging clocks we constructed have already reached saturation within the framework of linear models, a larger size of training data, in the hundreds or even thousands, may further enable building AI models to further increase accuracy. Furthermore, we generated single-cell type transcriptome clocks for each cell type and placed SCs to an even younger cross cell type median transcriptome age range of 65.01 to 75.83 years (Fig. 3C), much younger than their chronological ages. This also suggests a stronger slow aging of SCs at the single-cell transcriptome level than at the cell type composition level. It should be noted that all the age predictors we used (PLSR and EN) are linear models. Therefore, the existing models are sufficient to capture the linear age-related changes without necessarily the full age range. It is based on only these linear models we find the SCs have a significantly slower aging rate compared with expected linear changes during natural aging. In the future, it would be interesting to investigate the nonlinear changes during aging and in SCs, which will entail more samples from all age ranges for building nonlinear models. It would also be interesting to test the clock on progeroid diseases like Hutchinson-Gilford progeria syndrome.

To identify the cell types and molecular events underlying the slow aging of SCs, we used the clock models to generate the age-matched expected profiles for the SCs and then compared them to the real values in SCs in terms of cell types and molecular features. We found that, at the cell type level, SCs display an increase in naive CD8[+] T cells, together with a decrease in cytotoxic CD8[+] T cells and memory CD4[+] T cells. At the single-cell level in the HI-LR cell types, SCs contain more cells with high ribosome level, which is strongly negatively associated to the expression of inflammatory cytokine genes. Using a BN inference model, we inferred that high ribosome gene expression levels potentially contribute to a low inflammation state and delayed aging or slow aging of SCs in HI-LR cell types. Such a causal relationship was observed in cultured monocytes experimentally inhibiting ribosomal activity or translation induced the expression of inflammatory cytokines, such as IL-6 and IL-8. A recent study has found that these cytokines, generated by macrophages (tissue-infiltrated monocytes) if not rapidly dampened, can recruit cytotoxic CD8[+] T cells (*33*), which initiate killing of the senescent cells. It would be interesting to see whether the reduced level of cytotoxic CD8[+] T cells in SCs could be a further consequence of reduced inflammatory cytokine expression in the HI-LR cell types.

Ribosomopathy in humans results in Diamond Blackfan syndrome, which encompasses many age-related degenerative symptoms, including glaucoma and cataracts, at least through the G_0-G_1 cell cycle arrest (*34*). In yeast, ribosomal mutations result in shortened replicative life span (*35*). In contrast, improving translation efficiency by specific ribosome variant increases life span in multiple species (*36*). Our previous study has found ribosome gene expression being the most significant pathway positively correlated with life spans among mice subjected to different lifestyles (*37*). Here, we identified that single-cell ribosomal gene expression levels decrease with age in nearly all cell types and are significantly delayed in the HI-LR cell types in SCs (Fig. 4D). We also found that in the HI-LR cells, a further decrease in their basally low ribosomal gene expression could lead to enhanced inflammatory cytokine expression. Since ribosomal levels are a robust measure of growth rate for single cells (*33*, *38*) and enhanced inflammatory cytokine is a senescence-associated secretory phenotype (SASP) (*39*) accelerating aging rates in human (*8*), the high proportion of cells in high ribosome states of the HI-LR cell types in the blood from SCs and the ribosome/translation against inflammation balance identified by our single-cell aging clocks reveal an unexpected aging hallmark at the single-cell level and fuel the philosophical assertation "When you are not growing, you are not living."

MATERIALS AND METHODS

Ethics approval and participant consent

Human peripheral blood samples were collected from Shanghai East Hospital with informed consent and were approved by the hospital ethics committee in accordance with the Helsinki Declaration (registered number of approval: 2019.032).

Human blood sample collection and storage

Peripheral venous blood samples were collected at the Shanghai East Hospital Health Peer Center. Clinical parameters and medical histories were available for all volunteers. All volunteers were evaluated as generally healthy at the time of blood collection based on evaluation of their medical history and assessment of their vital signs. In our present study, 17 healthy volunteers were enrolled in scRNA-seq.

Isolation of PBMCs and monocyte

PBMCs were isolated from human peripheral blood samples as previously described (*40*). In brief, 5 ml of fresh peripheral blood was drawn per individual, mixed with 5 ml of phosphate-buffered saline (PBS), and then gently layered over 5 ml of Ficoll (GIBCO, USA) in a Falcon tube. Samples were centrifuged at 400*g* for 30 min without brake. After centrifugation, PBMCs in the white layer were carefully transferred to a new tube. After being washed with PBS twice, samples were then centrifuged at 400*g* for 10 min, and the pellets were resuspended with cryopreservation medium [1 ml of fetal bovine serum supplemented with 120 μl of dimethyl sulfoxide (DMSO)] for long-term storage in liquid nitrogen.

Human peripheral blood primary monocytes were isolated using the EasySep Human Monocyte Isolation Kit (STEMCELL Technologies, USA). PBMCs (5×10^7) were resuspended in 2 ml of isolation buffer (PBS containing 2% fetal bovine serum, 1 mM EDTA). Cells were incubated with 100 μl of isolation cocktail for 10 min under 4°C, and then 100 μl of magnetic particles was added, top up to

2.5 ml with isolation buffer, and incubated for 2.5 min at room temperature. Finally, the enriched cell suspension was harvested and centrifuged at 400g for 10 min.

10X genomics scRNA-seq library preparation

To perform scRNA-seq, thawed PBMCs were captured using the 10X Chromium System (10X Genomics, USA) according to the manufacturer's protocol. Cells were loaded onto the Chromium Controller Instrument to generate gel bead-in-emulsions (GEM). Afterward, barcoding and complementary DNA (cDNA) libraries were prepared using the Single-Cell 3′ Library and Gel Bead Kit (catalog no. 120237, 10X Genomics). GEM–reverse transcription (RT) was performed in a Mastercycler Nexus Thermal cycler (Eppendorf), and then GEMs were broken. cDNA was cleaned up with Dynabeads MyOne Silane Beads (Thermo Fisher Scientific, USA) and amplified. Amplified cDNA product was then cleaned up using the SPRIselect Reagent Kit (0.6× SPRI, Beckman Coulter). Indexed sequencing libraries were constructed using the Chromium Single-Cell 3′ Library Kit and subsequently sequenced using Illumina HiSeq PE150 at Novogene (Beijing, China) with average sequencing depth of >15 million reads per cell.

scRNA-seq data processing

Reads for each gene from the 17 scRNA-seq samples were quantified using the Cell Ranger Single-Cell Software Suite (version 3.0.0, 10X Genomics) against the GRCh38 human reference genome. Independent health cohorts of Chinese were downloaded from Beijing Institute of Genomics (BIG) Data Center: CYCT (41), HRA000150; WHCT (24), HRA000069. JOCT dataset was obtained from the authors of the original publication, and the SC dataset was downloaded from https://humandbs.biosciencedbc.jp/en/hum0229-v1 (21).

SCANPY (42) (single-cell analysis python toolkit, version 1.7.1) was used to read a combined gene-barcode matrix of these samples. Quality of cells was assessed based on three metrics step by step: (i) Genes expressed in <3 cells were removed. (ii) Cells with <200 or >5000 genes detected were removed. (iii) Cells with >15% mitochondrial gene counts were removed.

Batch effect correction and cell subtype annotations

To integrate cells into a shared space from different datasets for unsupervised clustering, we used the harmony algorithm to do batch effect correction (43). First, total count matrix was normalized by library size correction using default size factor 10,000. Highly variable genes (HVGs) were calculated by function pp.highly_variable_genes with default parameters. All ribosomal and mitochondrial genes were then removed from the list, which were identified as input to perform PCA analysis. Then, this PCA matrix was fed into pp.harmony-integrate function with default parameters, in which each sample was set as different batch. The resulting batch-corrected matrix was used to build nearest neighbor graph. Such nearest neighbor graph was then used to find clusters by Louvain algorithm with parameter "resolution" = 0.9 to identify the cell clusters. The cluster-specific marker genes were identified using the rank_genes_groups function by Wilcoxon rank-sum test. The top-ranked genes from the respective test statistic are regarded as marker genes. We annotated the clusters into eight major cell types based on the known marker genes. For gene expression

analysis, Combat (44) parametric adjustments were used for batch effect correction.

To identify subclusters within each major cell type, we performed a second round of clustering on T/NK, B, and myeloid cells separately. The procedure of the second round of clustering is the same as the first round on the HVGs chosen as described above, with resolution ranging from 0.1 to 0.9. Annotation of the resulting clusters to cell types was based on the known markers and top highly expressed genes compared with other cells.

Clock model generation

Partial least-squares regression (PLSR) clock model based on cell type compositions was established using the "plsr′ R packages with only the first three PLSR components to avoid overfitting. The leave-one-out (LOO) method (predictors were trained using all but one samples and then used to predict the age of the left-out samples) was applied to obtain the predicted age of each sample. Variable Influence on Projection (VIP) is a weighted sum of squares of the PLSR loadings that takes into account the amount of explained Y variance of each component. VIP scores give information about how the variables combine to form the quantitative relation between X and Y, thereby providing a better assessment of their relative importance in the model, and a VIP score ≥ 1.0 is considered as highly influential (45). VIP scores are calculated with the formula defined by Chong and Jun (46).

EN clock model was trained on all subtype compositions of 28 SE + GM samples with $\alpha = 5 \times 10^{-4}$ and L1 ratio = 0.08. Then, the cell types with high weight (|weight| > 50) in the model were considered to be important for age prediction.

We used the sum of gene expressions of all cells belonging to each cell type for each individual to train PLSR models as the pseudo-bulk transcriptome clock for the cell type.

Saturation analysis

To test the sufficiency of our sample size, we conducted a saturation analysis. We randomly selected one sample at a time to add to the training dataset, ranging from 5 to 28 samples. For each point, we repeated the process 100 times and calculated the average MAD and PCC.

Cell type annotation transfer by TOSICA

To transfer cell type annotations on additional datasets including GM old samples and Cord samples (GSE157007), COVID-19 (GSE158055), and SLE (GSE174188), we used the cell type annotation transfer tool TOSICA that we recently developed (26), which can achieve interpretable, accurate, and batch immune cell type transfer from reference dataset to query dataset. Here, we used 17 SEs as reference dataset to train a REACTOME pathway masked TOSICA model and applied the model to predict the cell type of each cell in the disease datasets.

Delayed age-up and age-down cell types

For PLSR-based procedures, the predict cell type compositions at 110 years were generated by calculating the solution of the reversed PLSR clock models at 110 years. $[\text{Pre\_compositions}]F_{n\times1} = 110 \times [\text{PLS1} + \text{PLS2}]^{-1}_{1\times n}$, where n is the 29 cell types used in the model. For LR model–based procedures, the delayed age-up and delayed age-down cell types are identified

by regressing the cell type compositions of each individual to age using the function lm(Composition~Age) in R for each gene with $P < 0.05$ (corrected by Benjamini-Hochberg adjustment). Then, the predict gene expression levels at 110 years were considered as the expected compositions, and z score of the expected value on the distribution of real composition levels in SCs was used to determine whether the gene expression was significantly delayed with $P < 0.05$ (corrected by Benjamini-Hochberg adjustment).

Delayed age-up and age-down genes, Gene Ontology terms, and TF enrichment analysis

For PLSR-based procedures, the predict gene expression levels at 110 years were generated by calculating the solution of the reversed PLSR clock models at 110 years. $[\text{Pre\_expression levels}]_{n \times 1} = 110 \times [\text{PLS1} + \text{PLS2}]^{-1}_{1 \times n}$, where n is the genes with VIP > 1. For LR model–based procedures, the delayed age-up and delayed age-down genes in each cell type are identified by regressing the median gene expression levels of these genes in each cell type of each individual to age using the function lm(Expression~Age) in R for each gene with $P < 0.05$. Then, the predict gene expression levels at 110y were considered as the expected expression, and z score of the expected value on the distribution of real expression levels in SCs was used to determine whether the gene expression was significantly delayed with $P < 0.05$.

Enriched KEGG pathway was determined for each group of cells using R package clusterProfiler (47) with default parameters. Annotation Dbi R package "org.Hs.eg.db" was used to map gene identifiers. Terms enriched in at least two cell types were visualized with the ggplot2 R package as dot plot. Metascape (http://metascape.org/gp/index.html#/main/step1) was used for TF enrichment analysis.

Inflammatory score

For a given cell or cell subset X, the inflammatory score was computed as the sum of all unique molecular identifier (UMI) for all the genes of "HALLMARK_INFLAMMATORY_RESPONSE" label from MsigDB (48), expressed in X, divided by the sum of all UMI expressed by X (49).

Cell-cell communication analysis

Cell-cell communication was predicted based on the scRNA-seq data by using CellPhoneDB software (version 1.1.0) (28). For each individual, only receptors and ligands expressed in at least 10% of cells of a given type were further analyzed. The significance of age-related difference was determined by RCC. The delayed age-up and down pairs were determined as delayed genes based on average expression of each ligand-receptor pair across different cell types, and only the top 100 pairs were used for visualization by LRPlot function from R package iTalk (50).

Age correction of AgeDiff and predicted age

To study chronological age-independent associations of AgeDiff (the difference between predicted and actual age based on PLS/EN model), we corrected AgeDiff for age by fitting to a polynomial model to age as follows, with span 0.5 and degree 2:

$$\text{AgeDiff} = \text{PredictedAge} - \text{Age}$$

$$\text{cAgeDiff} = \text{AgeDiff} - \text{loess(AgeDiff} \sim \text{Age)}$$

The following formula is used to correct the predicted age for the bias against age:

$$\text{cPredictedAge} = \text{PredictedAge} - \text{loess(AgeDiff} \sim \text{Age)}$$

BN inference

We used the WinMine package (https://www.microsoft.com/en-us/research/publication/the-winmine-toolkit-2) to calculate joint conditional probability and build the preliminary potential BNs. The BN analysis method is as described previously (51). Here, BN is inferred on the mean ribosome protein gene expression level [derived from scanpy.pp.calculate_qc_metrics() function], inflammation score, age, and age-corrected AgeDiff of each single cell. To avoid any signal loss by batch effect correction among four different datasets, the precorrection SE gene counts given by SCANPY were used. Only edges with PCC slope $P < 0.05$ are kept.

Translational inhibition assay

CHX (100 μg/ml final concentration) or equal volume of vehicle solvent (DMSO) was added to the THP-1 and human blood primary monocyte cells 15, 30, 60, 180, and 360 min before harvesting. Total RNA was extracted from cells using TRIzol Universal reagent (Tiangen, Beijing) according to the manufacturer's protocol. cDNA was synthesized using HiScript III RT SuperMix for qPCR (+gDNA wiper) (Vazyme, Nanjing), gene expression level was quantified by reverse transcription polymerase chain reaction (RT-PCR) using ChamQ Universal SYBR qPCR Master Mix (Vazyme, Nanjing), and relative gene expression levels were calculated using the $2^{-\triangle\triangle Ct}$ method. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as the endogenous control. The following primers were used for the quantitative PCR: GAPDH, 5'-TTGAGGTCAATGAAGGGGTC-3' (forward) and 5'-GAAGG TGAAGGTCGGAGTCA-3' (reverse); IL-6, 5'-ACTCACCTCTT CAGAACGAATTG-3' (forward) and 5'-CCATCTTTGGAA GGTTCAGGTTG-3' (reverse); IL-8, 5'-GTTTTTGAAGAGGGCT GAGAATTC-3' (forward) and 5'-CCCTACAACAGACCCACA CAATAC-3' (reverse).

REFERENCES AND NOTES

1. X. Xia, Y. Wang, Z. Yu, J. Chen, J.-D. J. Han, Assessing the rate of aging to monitor aging itself. *Ageing Res. Rev.* **69**, 101350 (2021).
2. M. J. Peters, R. Joehanes, L. C. Pilling, C. Schurmann, K. N. Conneely, J. Powell, E. Reinmaa, G. L. Sutphin, A. Zhernakova, K. Schramm, Y. A. Wilson, S. Kobes, T. Tukiainen; NABEC/UKBEC Consortium, M. A. Nalls, D. G. Hernandez, M. R. Cookson, R. J. Gibbs, J. Hardy, A. Ramasamy, A. B. Zonderman, A. Dillman, B. Traynor, C. Smith, D. L. Longo, D. Trabzuni, J. Troncoso, M. van der Brug, M. E. Weale, R. O'Brien, R. Johnson, R. Walker, R. H. Zielke, S. Arepalli, M. Ryten, A. B. Singleton, Y. F. Ramos, H. H. H. Göring, M. Fornage, Y. Liu, S. A. Gharib, B. E. Stranger, P. L. de Jager, A. Aviv, D. Levy, J. M. Murabito, P. J. Munson, T. Huan, A. Hofman, A. G. Uitterlinden, F. Rivadeneira, J. van Rooij, L. Stolk, L. Broer, M. M. P. J. Verbiest, M. Jhamai, P. Arp, A. Metspalu, L. Tserel, L. Milani, N. J. Samani, P. Peterson, S. Kasela, V. Codd, A. Peters, C. K. Ward-Caviness, C. Herder, M. Waldenberger,

M. Roden, P. Singmann, S. Zeilinger, T. Illig, G. Homuth, H. J. Grabe, H. Völzke, L. Steil, T. Kocher, A. Murray, D. Melzer, H. Yaghootkar, S. Bandinelli, E. K. Moses, J. W. Kent, J. E. Curran, M. P. Johnson, S. Williams-Blangero, H. J. Westra, A. F. McRae, J. A. Smith, S. L. R. Kardia, I. Hovatta, M. Perola, S. Ripatti, V. Salomaa, A. K. Henders, N. G. Martin, A. K. Smith, D. Mehta, E. B. Binder, K. M. Nylocks, E. M. Kennedy, T. Klengel, J. Ding, A. M. Suchy-Dicey, D. A. Enquobahrie, J. Brody, J. I. Rotter, Y. D. I. Chen, J. Houwing-Duistermaat, M. Kloppenburg, P. E. Slagboom, Q. Helmer, W. den Hollander, S. Bean, T. Raj, N. Bakhshi, Q. P. Wang, L. J. Oyston, B. M. Psaty, R. P. Tracy, G. W. Montgomery, S. T. Turner, J. Blangero, I. Meulenbelt, K. J. Ressler, J. Yang, L. Franke, J. Kettunen, P. M. Visscher, G. G. Neely, R. Korstanje, R. L. Hanson, H. Prokisch, L. Ferrucci, T. Esko, A. Teumer, J. B. J. van Meurs, A. D. Johnson, The transcriptional landscape of age in human peripheral blood. *Nat. Commun.* **6**, 8570 (2015).

3. B. Lehallier, D. Gate, N. Schaum, T. Nanasi, S. E. Lee, H. Yousef, P. Moran Losada, D. Berdnik, A. Keller, J. Verghese, S. Sathyan, C. Franceschi, S. Milman, N. Barzilai, T. Wyss-Coray, Undulating changes in human plasma proteome profiles across the lifespan. *Nat. Med.* **25**, 1843–1850 (2019).

4. G. Hannum, J. Guinney, L. Zhao, L. Zhang, G. Hughes, S. V. Sadda, B. Klotzle, M. Bibikova, J. B. Fan, Y. Gao, R. Deconde, M. Chen, I. Rajapakse, S. Friend, T. Ideker, K. Zhang, Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* **49**, 359–367 (2013).

5. S. Horvath, DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, 3156 (2013).

6. N. Sayed, Y. Huang, K. Nguyen, Z. Krejciova-Rajaniemi, A. P. Grawe, T. Gao, R. Tibshirani, T. Hastie, A. Alpert, L. Cui, T. Kuznetsova, Y. Rosenberg-Hasson, R. Ostan, D. Monti, B. Lehallier, S. S. Shen-Orr, H. T. Maecker, C. L. Dekker, T. Wyss-Coray, C. Franceschi, V. Jojic, F. Haddad, J. G. Montoya, J. C. Wu, M. M. Davis, D. Furman, An inflammatory aging clock (iAge) based on deep learning tracks multimorbidity, immunosenescence, frailty and cardiovascular aging. *Nat. Aging* **1**, 598–615 (2021).

7. W. Chen, W. Qian, G. Wu, W. Chen, B. Xian, X. Chen, Y. Cao, C. D. Green, F. Zhao, K. Tang, J. D. J. Han, Three-dimensional human facial morphologies as robust aging markers. *Cell Res.* **25**, 574–587 (2015).

8. X. Xia, X. Chen, G. Wu, F. Li, Y. Wang, Y. Chen, M. Chen, X. Wang, W. Chen, B. Xian, W. Chen, Y. Cao, C. Xu, W. Gong, G. Chen, D. Cai, W. Wei, Y. Yan, K. Liu, N. Qiao, X. Zhao, J. Jia, W. Wang, B. K. Kennedy, K. Zhang, C. V. Cannistraci, Y. Zhou, J. D. J. Han, Three-dimensional facial-image analysis to predict heterogeneity of the human ageing rate and the impact of lifestyle. *Nat. Metab.* **2**, 946–957 (2020).

9. S. Wang, Y. Zheng, J. Li, Y. Yu, W. Zhang, M. Song, Z. Liu, Z. Min, H. Hu, Y. Jing, X. He, L. Sun, L. Ma, C. R. Esteban, P. Chan, J. Qiao, Q. Zhou, J. C. Izpisua Belmonte, J. Qu, F. Tang, G. H. Liu, Single-cell transcriptomic atlas of primate ovarian aging. *Cell* **180**, 585–600.e19 (2020).

10. M. S. Shin, K. Yim, K. Moon, H. J. Park, S. Mohanty, J. W. Kim, R. R. Montgomery, A. C. Shaw, S. Krishnaswamy, I. Kang, Dissecting alterations in human CD8⁺ T cells with aging by high-dimensional single cell mass cytometry. *Clin. Immunol.* **200**, 24–30 (2019).

11. Z. Shi, Y. Geng, J. Liu, H. Zhang, L. Zhou, Q. Lin, J. Yu, K. Zhang, J. Liu, X. Gao, C. Zhang, Y. Yao, C. Zhang, Y. E. Sun, Single-cell transcriptomics reveals gene signatures and alterations associated with aging in distinct neural stem/progenitor cell subpopulations. *Protein Cell* **9**, 351–364 (2018).

12. H. Yousef, C. J. Czupalla, D. Lee, M. B. Chen, A. N. Burke, K. A. Zera, J. Zandstra, E. Berber, B. Lehallier, V. Mathur, R. V. Nair, L. N. Bonanno, A. C. Yang, T. Peterson, H. Hadeiba, T. Merkel, J. Körbelin, M. Schwaninger, M. S. Buckwalter, S. R. Quake, E. C. Butcher, T. Wyss-Coray, Aged blood impairs hippocampal neural precursor activity and activates microglia via brain endothelial cell VCAM1. *Nat. Med.* **25**, 988–1000 (2019).

13. J. C. Kimmel, L. Penland, N. D. Rubinstein, D. G. Hendrickson, D. R. Kelley, A. Z. Rosenthal, Murine single-cell RNA-seq reveals cell-identity- and tissue-specific trajectories of aging. *Genome Res.* **29**, 2088–2103 (2019).

14. Y. Elyahu, I. Hekselman, I. Eizenberg-Magar, O. Berner, I. Strominger, M. Schiller, K. Mittal, A. Nemirovsky, E. Eremenko, A. Vital, E. Simonovsky, V. Chalifa-Caspi, N. Friedman, E. Yeger-Lotem, A. Monsonego, Aging promotes reorganization of the CD4 T cell landscape toward extreme regulatory and effector phenotypes. *Sci. Adv.* **5**, eaaw8330 (2019).

15. C. P. Martinez-Jimenez, N. Eling, H. C. Chen, C. A. Vallejos, A. A. Kolodziejczyk, F. Connor, L. Stojic, T. F. Rayner, M. J. T. Stubbington, S. A. Teichmann, M. de la Roche, J. C. Marioni, D. T. Odom, Aging increases cell-to-cell transcriptional variability upon immune stimulation. *Science* **355**, 1433–1436 (2017).

16. I. Angelidis, L. M. Simon, I. E. Fernandez, M. Strunz, C. H. Mayr, F. R. Greiffo, G. Tsitsiridis, M. Ansari, E. Graf, T. M. Strom, M. Nagendran, T. Desai, O. Eickelberg, M. Mann, F. J. Theis, H. B. Schiller, An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics. *Nat. Commun.* **10**, 963 (2019).

17. M. Ximerakis, S. L. Lipnick, B. T. Innes, S. K. Simmons, X. Adiconis, D. Dionne, B. A. Mayweather, L. Nguyen, Z. Niziolek, C. Ozek, V. L. Butty, R. Isserlin, S. M. Buchanan, S. S. Levine, A. Regev, G. D. Bader, J. Z. Levin, L. L. Rubin, Single-cell transcriptomic profiling of the aging mouse brain. *Nat. Neurosci.* **22**, 1696–1708 (2019).

18. A. Trapp, C. Kerepesi, V. Gladyshev, Profiling epigenetic age in single cells. *Innov. Aging* **5**, 673 (2021).

19. M. T. Buckley, E. D. Sun, B. M. George, L. Liu, N. Schaum, L. Xu, J. M. Reyes, M. A. Goodell, I. L. Weissman, T. Wyss-Coray, T. A. Rando, A. Brunet, Cell-type-specific aging clocks to quantify aging and rejuvenation in neurogenic regions of the brain. *Nat. Aging* **3**, 121–137 (2023).

20. M. G. P. van der Wijst, H. Brugge, D. H. de Vries, P. Deelen, M. A. Swertz; Life Lines Cohort Study; BIOS Consortium, L. Franke, Single-cell RNA sequencing identifies celltype-specific cis-eQTLs and co-expression QTLs. *Nat. Genet.* **50**, 493–497 (2018).

21. K. Hashimoto, T. Kouno, T. Ikawa, N. Hayatsu, Y. Miyajima, H. Yabukami, T. Terooatea, T. Sasaki, T. Suzuki, M. Valentine, G. Pascarella, Y. Okazaki, H. Suzuki, J. W. Shin, A. Minoda, I. Taniuchi, H. Okano, Y. Arai, N. Hirose, P. Carninci, Single-cell transcriptomics reveals expansion of cytotoxic CD4 T cells in supercentenarians. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 24242–24251 (2019).

22. O. J. Luo, W. Lei, G. Zhu, Z. Ren, Y. Xu, C. Xiao, H. Zhang, J. Cai, Z. Luo, L. Gao, J. Su, L. Tang, W. Guo, H. Su, Z. J. Zhang, E. F. Fang, Y. Ruan, S. X. Leng, Z. Ju, H. Lou, J. Gao, N. Peng, J. Chen, Z. Bao, F. Liu, G. Chen, Multidimensional single-cell analysis of human peripheral blood reveals characteristic features of the immune system landscape in aging and frailty. *Nat. Aging* **2**, 348–364 (2022).

23. J. Y. Zhang, X. M. Wang, X. Xing, Z. Xu, C. Zhang, J. W. Song, X. Fan, P. Xia, J. L. Fu, S. Y. Wang, R. N. Xu, X. P. Dai, L. Shi, L. Huang, T. J. Jiang, M. Shi, Y. Zhang, A. Zumla, M. Maeurer, F. Bai, F. S. Wang, Single-cell landscape of immunological responses in patients with COVID-19. *Nat. Immunol.* **21**, 1107–1118 (2020).

24. W. Wen, W. Su, H. Tang, W. le, X. Zhang, Y. Zheng, X. Liu, L. Xie, J. Li, Y. Ye, L. Dong, X. Cui, Y. Miao, D. Wang, J. Dong, C. Xiao, W. Chen, H. Wang, Immune cell profiling of COVID-19 patients in the recovery stage by single-cell sequencing. *Cell Discov.* **6**, 31 (2020).

25. M. Farres, S. Platikanov, S. Tsakovski, R. Tauler, Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation. *J. Chemometr.* **29**, 528–536 (2015).

26. J. Chen, H. Xu, W. Tao, Z. Chen, Y. Zhao, J. D. J. Han, Transformer for one stop interpretable cell type annotation. *Nat. Commun.* **14**, 223 (2023).

27. C. Lopez-Otin, M. A. Blasco, L. Partridge, M. Serrano, G. Kroemer, The hallmarks of aging. *Cell* **153**, 1194–1217 (2013).

28. R. Vento-Tormo, M. Efremova, R. A. Botting, M. Y. Turco, M. Vento-Tormo, K. B. Meyer, J. E. Park, E. Stephenson, K. Polański, A. Goncalves, L. Gardner, S. Holmqvist, J. Henriksson, A. Zou, A. M. Sharkey, B. Millar, B. Innes, L. Wood, A. Wilbrey-Clark, R. P. Payne, M. A. Ivarsson, S. Lisgo, A. Filby, D. H. Rowitch, J. N. Bulmer, G. J. Wright, M. J. T. Stubbington, M. Haniffa, A. Moffett, S. A. Teichmann, Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* **563**, 347–353 (2018).

29. X. Ren, W. Wen, X. Fan, W. Hou, B. Su, P. Cai, J. Li, Y. Liu, F. Tang, F. Zhang, Y. Yang, J. He, W. Ma, J. He, P. Wang, Q. Cao, F. Chen, Y. Chen, X. Cheng, G. Deng, X. Deng, W. Ding, Y. Feng, R. Gan, C. Guo, W. Guo, S. He, C. Jiang, J. Liang, Y. M. Li, J. Lin, Y. Ling, H. Liu, J. Liu, N. Liu, S. Q. Liu, M. Luo, Q. Ma, Q. Song, W. Sun, G. X. Wang, F. Wang, Y. Wang, X. Wen, Q. Wu, G. Xu, X. Xie, X. Xiong, X. Xing, H. Xu, C. Yin, D. Yu, K. Yu, J. Yuan, B. Zhang, P. Zhang, T. Zhang, J. Zhao, P. Zhao, J. Zhou, W. Zhou, S. Zhong, X. Zhong, S. Zhang, L. Zhu, P. Zhu, B. Zou, J. Zou, Z. Zuo, F. Bai, X. Huang, P. Zhou, Q. Jiang, Z. Huang, J. X. Bei, L. Wei, X. W. Bian, X. Liu, T. Cheng, X. Li, P. Zhao, F. S. Wang, H. Wang, B. Su, Z. Zhang, K. Qu, X. Wang, J. Chen, R. Jin, Z. Zhang, COVID-19 immune features revealed by a large-scale single-cell transcriptome atlas. *Cell* **184**, 5838 (2021).

30. M. M. Markiewski, S. K. Vadrevu, S. K. Sharma, N. K. Chintala, S. Ghouse, J. H. Cho, D. P. Fairlie, Y. Paterson, A. Astrinidis, M. Karbowniczek, The ribosomal protein S19 suppresses antitumor immune responses via the complement C5a receptor 1. *J. Immunol.* **198**, 2989–2999 (2017).

31. T. Schneider-Poetsch, J. Ju, D. E. Eyler, Y. Dang, S. Bhat, W. C. Merrick, R. Green, B. Shen, J. O. Liu, Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nat. Chem. Biol.* **6**, 209–217 (2010).

32. D. Cai, J.-D. J. Han, Aging-associated lncRNAs are evolutionarily conserved and participate in NFκB signaling. *Nat. Aging* **1**, 438–453 (2021).

33. I. Sturmlechner, C. Zhang, C. C. Sine, E. J. van Deursen, K. B. Jeganathan, N. Hamada, J. Grasic, D. Friedman, J. T. Stutchman, I. Can, M. Hamada, D. Y. Lim, J. H. Lee, T. Ordog, R. M. Laberge, V. Shapiro, D. J. Baker, H. Li, J. M. van Deursen, p21 produces a bioactive secretome that places stressed cells under immunosurveillance. *Science* **374**, eabb3420 (2021).

34. K. R. Kampen, S. O. Sulima, S. Vereecke, K. De Keersmaecker, Hallmarks of ribosomopathies. *Nucleic Acids Res.* **48**, 1013–1028 (2020).

35. A. W. MacInnes, The role of the ribosome in the regulation of longevity and lifespan extension. *WIREs RNA* **7**, 198–212 (2016).

36. V. E. Martinez-Miguel, C. Lujan, T. Espie-Caullet, D. Martinez-Martinez, S. Moore, C. Backes, S. Gonzalez, E. R. Galimov, A. E. X. Brown, M. Halic, K. Tomita, C. Rallis, T. von der Haar,

F. Cabreiro, I. Bjedov, Increased fidelity of protein synthesis extends lifespan. *Cell Metab.* **33**, 2288–2300.e12 (2021).

37. B. Zhou, L. Yang, S. Li, J. Huang, H. Chen, L. Hou, J. Wang, C. D. Green, Z. Yan, X. Huang, M. Kaeberlein, L. Zhu, H. Xiao, Y. Liu, J. D. Han, Midlife gene expressions identify modulators of aging through dietary interventions. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E1201–E1209 (2012).

38. M. Jung, S. G. Jin, X. Zhang, W. Xiong, G. Gogoshin, A. S. Rodin, G. P. Pfeifer, Longitudinal epigenetic and gene expression profiles analyzed by three-component analysis reveal down-regulation of genes involved in protein translation in human aging. *Nucleic Acids Res.* **43**, e100 (2015).

39. J. P. Coppe, P. Y. Desprez, A. Krtolica, J. Campisi, The senescence-associated secretory phenotype: The dark side of tumor suppression. *Annu. Rev. Pathol.* **5**, 99–118 (2010).

40. W. Wen, J. P. Zhang, J. Chen, C. Arakaki, X. Li, D. Baylink, G. D. Botimer, J. Xu, W. Yuan, T. Cheng, X. B. Zhang, Generation of integration-free induced pluripotent stem cells from human peripheral blood mononuclear cells using episomal vectors. *J. Vis. Exp.*, 55091 (2017).

41. J. Tanevski, T. Nguyen, B. Truong, N. Karaiskos, M. E. Ahsen, X. Zhang, C. Shu, K. Xu, X. Liang, Y. Hu, H. V. V. Pham, L. Xiaomei, T. D. le, A. L. Tarca, G. Bhatti, R. Romero, N. Karathanasis, P. Loher, Y. Chen, Z. Ouyang, D. Mao, Y. Zhang, M. Zand, J. Ruan, C. Hafemeister, P. Qiu, D. Tran, T. Nguyen, A. Gabor, T. Yu, J. Guinney, E. Glaab, R. Krause, P. Banda; DREAM SCTC Consortium, G. Stolovitzky, N. Rajewsky, J. Saez-Rodriguez, P. Meyer, Gene selection for optimal prediction of cell position in tissues from single-cell transcriptomics data. *Life Sci. Alliance* **3**, e202000867 (2020).

42. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).

43. I. Korsunsky, N. Millard, J. Fan, K. Slowikowski, F. Zhang, K. Wei, Y. Baglaenko, M. Brenner, P. R. Loh, S. Raychaudhuri, Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).

44. W. E. Johnson, C. Li, A. Rabinovic, Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).

45. J. Xia, I. V. Sinelnikov, D. S. Wishart, MetATT: A web-based metabolomics tool for analyzing time-series and two-factor datasets. *Bioinformatics* **27**, 2455–2456 (2011).

46. I. G. Chong, C. H. Jun, Performance of some variable selection methods when multicollinearity is present. *Chemometr. Intell. Lab.* **78**, 103–112 (2005).

47. G. C. Yu, L.-G. Wang, Y. Y. Han, Q.-Y. He, clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).

48. A. Liberzon, C. Birger, H. Thorvaldsdóttir, M. Ghandi, J. P. Mesirov, P. Tamayo, The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).

49. H. Leins, M. Mulaw, K. Eiwen, V. Sakk, Y. Liang, M. Denkinger, H. Geiger, R. Schirmbeck, Aged murine hematopoietic stem cells drive aging-associated immune remodeling. *Blood* **132**, 565–576 (2018).

50. Y. Wang, R. Wang, S. Zhang, C. Jiang, G. Han, M. Wang, J. Ajani, A. Futreal, L. Wang, iTALK: An R package to characterize and illustrate intercellular communication. bioRxiv 507871 [Preprint]. 4 January 2019. https://doi.org/10.1101/507871.

51. H. Yu, S. S. Zhu, B. Zhou, H. L. Xue, J.-D. J. Han, Inferring causal relationships among different histone modifications and gene expression. *Genome Res.* **18**, 1324–1314 (2008).